1. SIGNALS AND INFORMATION AND CLASSIFICATION OF SIGNALS

1.0 TERMINOLOGY

Some preliminary definitions are required in order to be able to tackle the theory that is related to signals. Signal and Information theory use, indeed, a large set of words that often have a different meaning when those would be used in current day-by-day language.

Signal: The word signal originates from the Latin word *signum*, which means an object, a sign, a token, sometimes a gesture. The understanding that is associated with the word signal, therefore, is very old and dates back to the prehistory. Electrical signals, however, appear in engineering only since the 19th century[1]. *Every mathematical function, of which one of the variables is time, can be considered as a signal*. This definition, however, seems to be too wide. The acoustical sound from the dial tone in a telephone system indeed can be regarded as being a signal. The physical description of the sound will require a mathematical function of which the time will be a parameter. Planets moving around the sun can be described also by a mathematical function with time dependency, but are not regarded as being signals. Hence, the first attempt to propose a definition must be refined. *A mathematical function depending on the time will be called a process*. So, signals are processes, but the reverse is not necessarily true. The second less general attempt to define a signal could be: *a signal is a process that can propagate in a certain medium*. This definition confirms that a signal is a process, but indicates further that it is to be prepared by a source, that it will propagate in a physical medium further to reach a destination, where it will be received. An alternative definition that takes the previous concerns into consideration but that will introduce the new concept of *information* is of Frederic de Coulon: *A signal is the physical description of the information that it carries from a source towards a receiving destination*.

Message: A usually short communication transmitted by words, signals or all other means from a transmitter to one or more receivers. The shortest and hence *elementary message* is considered to be a *word*.

Channel: The set of tools that is used for the realization of the transmission of a signal.

Modulation: The conversion of a message into a signal ensuring better

[1] Invention of the telegraph (Morse, Cooke, Wheatstone, 1830-1840), the telephone (Bell, 1876) and the first radio communications (Popov, Marconi, 1895-1896). Electronics, from the beginning of the 20th century, allowed amplifying and detecting weak signals (Fleming, Lee de Forest, 1904-1907). This progress in technology forms the basis of modern signal theory and signal processing…

transmission of the latter in a certain physical medium or allowing multiple transmissions in the same physical medium will make use of *modulation*. The reverse action is called *demodulation*.

Code: The transformation of a message into a discrete signal in order to enhance the efficiency of the transmission is called *coding*. The list with agreed symbols is called a *code*.

Information: *Information* will be defined later, but it will be shown that it is a *measurable quantity*. When it is supposed that N realizations can be obtained (e.g. N=6 when using a dice) and that all realizations have equal probability (the dice is perfect; p = 1/N), then the information is given by the following log formula:

$$i = +\log_2 \frac{1}{p} = -\log_2 p = \log_2 N$$

1.1 ELEMENTS IN A COMMUNICATION SYSTEM

The most essential elements that are present in a communication system are depicted in Figure 1.1.

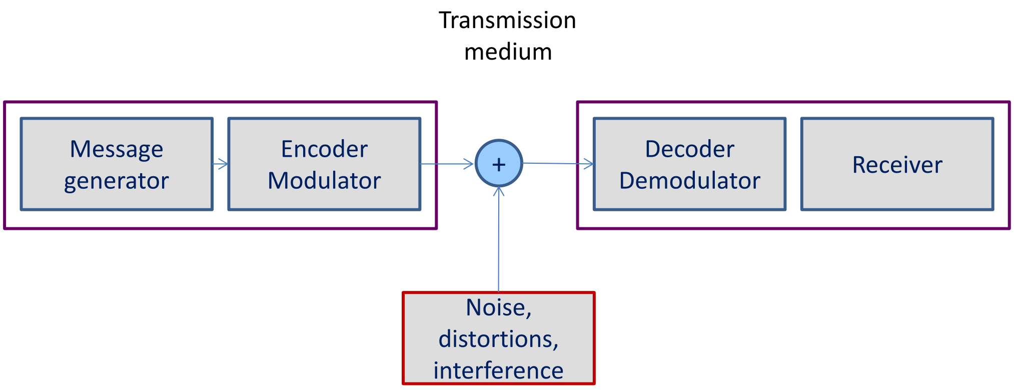

Figure 1.1: The necessary elements in a communication channel.

The *physical medium* (PHY in the OSI layer 1) used for the transmission can consist of electrical wires, optical fiber, air, etc. Such communication systems are trivial in day-by-day life (e.g. AM-FM radio broadcasting, telephone (POTS), file transfer between computers, internet access…). In Figure 1.2 as an example the connection is sketched between a PC in a home office to the Internet. Via a network the PC at home of the customer is connected to a server of an ISP (Internet Service Provider); e.g. using the cupper network of the POTS (Plain Old Telephony Service) in a P2P (Point-to-Point) DSL (Digital Subscriber Line) connection (ADSL2+, VDSL2 …).

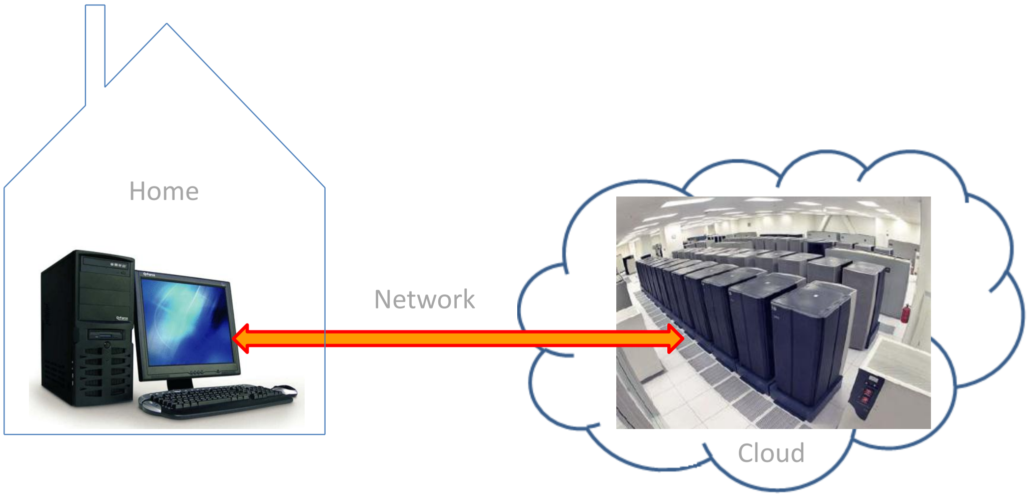

Figure 1.2: An example of a communication system. At PC at home is connected via a network to Internet.

Problem: Determine the transmission medium to be used in such a way that optimal transfer of messages will occur at a given cost. The implies that convertors must be developed to reduce the additive disturbances (noise, interference,...) to levels that are considered to be acceptable.

Two more general terms are important in communication:

Redundancy: A surplus of information; e.g. in transmitting the same message twice to overcome errors due to the noise in the channel. A source is redundant when it transmits symbols that are not completely independent of each other and that are not strictly necessary to convey the information. An example will clarify the principle: «Is it required to have a **u** following the character **q** in the English language to uniquely interpret a text»? One can state that redundancy in the English language is about 30%. This means that in a long text, up to one third of the characters can be considered as being superfluous for the transmission of the information. Example:

Yu shld b abl t read ths eventho sevrl ltrs r msng.[2]

{Check this out! In this example more the 33% of the characters have been removed}

Other languages can be less redundant. One of the least redundant languages is Kiswahili with around 10% characters that are superfluous. This is because the writing has been changed from Arabic to Roman characters end of the 19th century and because phonetic writing has been used. The sentence below illustrates this *(How are you? Not so well, because I am sick now)*.

[2] "You should be able to read this even though several letters are missing"

Habari gani? Mbaya mdogo, kwa sababu mimi ni mgonjwa sasa.

Ambiguity: uncertainty at the receiver about the message that has been transmitted. If the 1-to-1 mapping (each transmitted message will correspond to a unique received message) is failing, then eventually more than 1 meaning can be attributed to the received word. Such cases are e.g. present in Radar and Sonar signals.

1.2 CLASSIFICATION OF SIGNALS

1.2.1 DETERMINISTIC SIGNALS

The values of these signals can be obtained at any time instance using a mathematical model. In the simplest case this can be a formula, of which the time is one of the parameters. Hence, statistical processing or study for the analysis of the signal is not required. To ensure that a signal will be able to carry information, one will observe further that it is required that the latter should have at least some stochastic properties.

Example: the signal $s(t) = A_0 \cos(\omega_0 t + \varphi)$ where A_0, ω_0 and φ are constants, is purely deterministic. When the amplitude, pulsation and phase shift have known values, it is perfectly possible to compute the value that the signal will obtain at any freely chosen time instant t.

It has no real advantage to transmit such a signal over a transmission channel, since with the knowledge of the amplitude, phase and pulsation at any time instance the actual value can be computed, and hence regenerated easily. Such signal, as will be demonstrated further in chapter 4 does not convey any information.

1.2.2 STOCHASTIC SIGNALS

The values that the signal will gain at any time instance cannot be computed from a deterministic model and the behaviour, i.e. the successive values that a signal will take at any chosen time instance, is not known in advance, and hence seems to be unpredictable. In general such signals can be described using large sets of statistical observations. Hence, such signals will have to be treated in a statistical way.

1.2.3 PERIODICAL SIGNALS

This important class of signals ensures the basis for harmonic analysis.

They show high importance in theoretical studies. It will be shown that their properties can be applied to deterministic signals as well[3].

1.3 USE OF BASE FUNCTIONS

To represent even very complex signals, one can rely on the development (expansion) into series using simple signals as the basis. Such representations can be of interest to determine the properties of the transmission in linear systems. In case of non-linear systems, one should justify the series for each different case; e.g. using Volterra series.

In order to represent a signal in a series

$$s(t) = \sum_i a_i \; s_i(t)$$

it is necessary that the latter would converge to the correct values at any time instance. This can be analysed by the error term (the difference between the real signal at any time instant and the generated one from the time series.

∃ N ⊃- for ε>0 chosen arbitrarily:

$$\left\| s(t) - \sum_{i=1}^{M} a_i \, s_i(t) \right\| \leq \varepsilon \qquad \text{for } M>N$$

and since the definition for the Euclidian norm:

$$\| f \| \triangleq \sqrt{\int_a^b f^2(x) \; dx}$$

This results in:

$$\left\| s(t) - \sum_{i=1}^{M} a_i \, s_i(t) \right\| = \sqrt{\int_a^b \left[s(t) - \sum_{i=1}^{M} a_i s_i(t) \right]^2 dt}$$

[3] The reader is reminded that periodic signals exist that have a quasi-stochastic behaviour. A well-known class are the pseudorandom MLBS (Maximum Length Binary Signals), that can be generated using a shift register and an EXOR (Exclusive OR).

A useful, but not necessary condition is that the base functions should be orthonormal in an interval $[-k_1, +k_2]$ (e.g. $[-\infty, +\infty]$, $[0,T]$…); i.e. that:

$$\int_{-k_1}^{k_2} s_i(t) \,.\, s_j(t)\, dt = \delta_{ij}$$

This holds when the signals are of the real type. Else (complex signals), the complex conjugate must be inserted into the integrand. This condition allows easily retrieving the coefficients a_i.

Observe that:

$$\int_{-k_1}^{k_2} s(t)\, s_i(t)\, dt = \int_{-k_1}^{k_2} \left[\sum_j a_j\, s_j(t) \right] s_i(t)\, dt = \sum_j a_j \int_{-k_1}^{k_2} s_i(t)\, s_j(t)\, dt = a_i$$

And hence:

$$a_i = \int_{-k_1}^{k_2} s(t)\, s_i(t)\, dt$$

Often the base functions are relatively simple sets; e.g. $\cos(i\omega t)$, $\sin(i\omega t)$, or exponential functions, de Dirac distributions or the Heaviside function (unit step function).

1.4 INTEGRAL TRANSFORMATIONS

Using a transformation one can swap domains and convert time t to another variable α. The transformations are of the integral operation type.

$$F(\alpha) \triangleq \int_a^b g(\alpha,t) \,.\, f(t)\, dt$$

With $g(\alpha,t)$ the kernel of the transform.

Usually the **Fourier transform** is used. Then, the variables become:

$$\alpha = \omega$$
$$a = -\infty$$
$$b = +\infty$$

$$g(\alpha,t) = e^{-j\omega t}$$

Other example is the **Laplace transform**. Then, the variables are $\alpha=s$, $a=0$, $b=+\infty$, $g(\alpha,t) = e^{-st}$. The **Hilbert transform** is obtained for $\alpha=\tau$, $a=-\infty$, $b=+\infty$, $g(\alpha,t) = -[\pi(t-\tau)]^{-1}$, etc.

In the case the signals are time invariant and the systems linear, then for several transformations the convolution theorem holds; i.e. if:

$$\int_a^b g(\alpha,t)\, f_1(t)\, dt = F_1(\alpha)$$

And:

$$\int_a^b g(\alpha,t)\, f_2(t)\, dt = F_2(\alpha)$$

When the following equation is noted as:

$$\int_a^b f_1(\tau) \,.\, f_2(t-\tau)\, d\tau \equiv f_1(t) * f_2(t)$$

Then:

$$\int_a^b g(\alpha,t) \,.\, f_1(t) * f_2(t)\, dt = F_1(\alpha) \,.\, F_2(\alpha)$$

In other words the convolution in the time domain results in the product of the integral transforms of both signals.

This theorem is of high importance for analysis in the frequency domain.

If $f(t) \Leftrightarrow F(\omega)$, where $\Leftrightarrow$ means ‘has the spectrum…’, and for

$$f(t) \Leftrightarrow F(\omega)$$

$$g(t) \Leftrightarrow G(\omega)$$

the following expressions yield:

$$f(t)*g(t) \Leftrightarrow F(\omega).G(\omega)$$

$$f(t).g(t) \Leftrightarrow F(\omega)*G(\omega)/2\pi$$

Important theorem: independent from the domain of representation, the energy content of the signal remains unaltered (conservation principle). For the Fourier integral this is known as the theorem of **Parseval.**

If $F(\omega) = A(\omega).e^{-j\varphi(\omega)}$ represents the Fourier integral of f(t), then

$$\int_{-\infty}^{+\infty} \| f(t) \|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A^2(\omega)\, d\omega$$

Remark: note that the definition of Hilbert transformation is based on a convolution in the time domain. Indeed:

$$\mathcal{H}(g(t)) \triangleq -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{g(\tau)}{(t-\tau)} d\tau = g(t) * \left[\frac{-1}{\pi t}\right]$$

1.5 PROPERTIES OF STOCHASTIC SIGNALS

A stochastic signal is a process that develops in time and where at least for part probabilistic behaviour can be noted. Mathematically, a stochastic signal will depend on 2 variables: k and t.

$$\xi(k,t) = \xi^{(k)}(t)$$

k are values generated by the stochastic space and t is the time. At any arbitrarily time instance $t=t_1$, the value of the signal $\xi^{(k)}(t_1)$ will be generated, delivered by the stochastic space. Such attribution of values is called the *realization* of the process. Time instances can be either discrete for time instances $t_1, t_2, \ldots, t_n$ or continuous for arbitrarily values of t.

Example:

For a die, the stochastic space is discrete and accounts 6 values {1,2,3,4,5,6}. When the dice is thrown at time instance $t=t_1$, the generated number of eyes k will be regarded as the realization $\xi^{(k)}(t_1)$.

To study the statistical properties of stochastic signals one must rely on statistical techniques.

$P\{\xi(t_1)\leq x_1\}$: is the probability that the value of $\xi(t)$ at the time instance $t=t_1$ would be less or equal to x_1.

In general this probability will depend both on the chosen time instance t_1 and on the value x_1. Hence, the latter is a function of two variables. The function is called the *probability function* or the *distribution function*:

$F_1(x_1,t_1) = P\{\xi(t_1)\le x_1\}$ = the ratio of all realizations where $\xi(t_1)\le x_1$ to the total number of realizations.

From this distribution function the *density functions* are derived:

$$w_1(x_1,t_1) = \frac{\partial F_1(x_1,t_1)}{\partial x_1}$$

This function is called the *probability density function*, or p.d.f.

Following this principle, one obtains further the distribution function:

$F_n(x_1,x_2,\dots,x_n;t_1,t_2,\dots,t_n) = P\{\xi(t_1)\le x_1,\dots,\xi(t_n)\le x_n\}$

This is the probability that $\xi(t_1)\le x_1$ and $\xi(t_2)\le x_2$,... and $\xi(t_n)\le x_n$. Hence one can derive also:

$$w_n(x_1,x_2,\dots,x_n;t_1,t_2,\dots,t_n) = \frac{\partial^n F_n(x_1,x_2,\dots,x_n;t_1,t_2,\dots,t_n)}{\partial x_1\partial x_2\dots\partial x_n}$$

1.6 STATISTICAL MOMENTS

Ensemble averages of means are the statistical moments with respect to the stochastic space that is generating the signals. First order, second order and mixed moments will be treated in this study only. However, higher order moments can be of interest in certain application areas in science and engineering. Stochastic modelling of communication channels, for instance, invokes the moments of order 4.

1. **Mean**:

$$\overline{\xi(t_1)} \triangleq \int_{-\infty}^{+\infty} x_1\, w_1(x_1;t_1)\, dx_1 = \mu(t_1)$$

2. **Quadratic mean**:

$$\overline{\xi^2(t_1)} \triangleq \int_{-\infty}^{+\infty} x_1^2\, w_1(x_1;t_1)\, dx_1$$

3. **Variance**:

$$\sigma^2(t_1) \triangleq \overline{\left[\xi(t_1) - \overline{\xi(t_1)}\right]^2} = \overline{\xi(t_1)^2} - \left[\overline{\xi(t_1)}\right]^2$$

The variance yields the quadratic mean minus the square of the mean. <u>Proof</u>:

$$\sigma^2(t_1) \triangleq \overline{\left[\, \xi(t_1) \ - \ \overline{\xi(t_1)} \,\right]^2} \ = \ \overline{\xi(t_1)^2} \ + \ \overline{\left[\overline{\xi(t_1)}\right]^2} \ - \ 2 \left[\overline{\xi(t_1) \, . \, \overline{\xi(t_1)}}\right]$$

$$\sigma^2(t_1) \ = \ \overline{\xi(t_1)^2} \ + \ \overline{\xi(t_1)}^2 \ - \ 2 \ . \ \overline{\xi(t_1)} \ . \ \overline{\xi(t_1)}$$

and hence:

$$\sigma^2(t_1) \triangleq \overline{\left[\, \xi(t_1) \ - \ \overline{\xi(t_1)} \,\right]^2} \ = \ \overline{\xi(t_1)^2} \ - \ \left[\, \overline{\xi(t_1)} \,\right]^2$$

4. **Correlation function**:

$$B_{\xi\eta}(t_1,t_2) \ \triangleq \ \overline{\xi(t_1)\, \eta(t_2)} \ = \ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 \ y_2 \ w_2(x_1,y_2;t_1,t_2) \ dx_1 dy_2$$

This correlation function is called the *cross-correlation* to underline that ξ and η are different processes (signals). As the special case ($\eta=\xi$) the *autocorrelation* will be used:

$$B_{\xi\xi}(t_1,t_2) \ \triangleq \ \overline{\xi(t_1)\, \xi(t_2)} \ = \ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 \ x_2 \ w_2(x_1,x_2;t_1,t_2) \ dx_1 dx_2$$

5. **Covariance**:

$$K_{\xi\eta}(t_1,t_2) \ \triangleq \ \left[\xi(t_1) - \overline{\xi(t_1)}\right] . \left[\eta(t_2) - \overline{\eta(t_2)}\right] \ = \ B_{\xi\eta}(t_1,t_2) - \mu_\xi(t_1).\mu_\eta(t_2)$$

With also here, as special case, the *autocovariance* (when $\eta=\xi$). The proof of the equation is similar to the one for the variance. Hence, the *covariance* equals the *cross-correlation* minus the product of the *ensemble averages* for the processes ξ and η.

$$\begin{aligned}
\overline{\left[\xi(t_1) - \overline{\xi(t_1)}\right] \bullet \left[\eta(t_2) - \overline{\eta(t_2)}\right]} \ &= \ \overline{\xi(t_1).\eta(t_2) \ + \ \overline{\xi(t_1)}.\overline{\eta(t_2)} \ - \ \overline{\xi(t_1)}.\eta(t_2) \ - \ \xi(t_1).\overline{\eta(t_2)}} \\
&= \ \overline{\xi(t_1).\eta(t_2)} \ + \ \overline{\xi(t_1)}.\overline{\eta(t_2)} \ - \ 2\, \overline{\xi(t_1)}.\overline{\eta(t_2)} \\
&= \ B_{\xi\eta}(t_1,t_2) \ \ -\mu_\xi(t_1).\mu_\eta(t_2)
\end{aligned}$$

1.7 TIME AVERAGES OR REALIZATION MEANS

Often one does not have access to all the realizations that are produced by the stochastic space, but is one limited to a few of them, or even to one single version of the latter. Therefore, averages with respect to time are defined.

1. **Mean**:

$$\widetilde{\xi^{(k)}(t_0+t)} \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \xi^{(k)}(t_0+t)\,dt$$

The mean is independent of the origin of the time axis t_0, since the mean is calculated over all values in time $[-\infty,+\infty]$. This can be shown readily in substituting $t' = t_0 + t$. One obtains then:

$$\widetilde{\xi^{(k)}(t_0+t)} = \lim_{T\to\infty} \frac{1}{T} \int_{t_0-\frac{T}{2}}^{t_0+\frac{T}{2}} \xi^{(k)}(t')\,dt' = \widetilde{\xi^{(k)}(t)}$$

which is independent of the time origin t_0. The time average of mean of the realization represents the *D.C. component* of a signal.

2. **Quadratic mean**:

$$\widetilde{\left[\xi^{(k)}(t_0+t)\right]^2} \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \left[\xi^{(k)}(t_0+t)\right]^2 dt = \widetilde{\left[\xi^{(k)}(t)\right]^2}$$

Also this mean is independent of the time origin t_0 and it represents the *power* of the signal.

$$\sigma^2 = \widetilde{\xi^{(k)}(t_0)^2} - \left[\widetilde{\xi^{(k)}(t_0)}\right]^2$$

3. **Autocorrelation function**:

$$R_{\xi\xi}^{(k)}(t_1-t_2) \triangleq \widetilde{\xi^{(k)}(t_1+t)\,.\,\xi^{(k)}(t_2+t)}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \xi^{(k)}(t_1+t)\,.\,\xi^{(k)}(t_2+t)\,dt = R_{\xi\xi}^{(k)}(t_2-t_1)$$

The autocorrelation is independent of the time origin t_0, but depends only on the time interval $|t_1 - t_2|$. Indeed, define $t = -t_1 + t'$. Then one obtains:

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_1 - \frac{T}{2}}^{t_1 + \frac{T}{2}} \xi^{(k)}(t') \,.\, \xi^{(k)}(t_2 - t_1 + t')\, dt' = R_{\xi\xi}^{(k)}(t_2 - t_1)$$

With the substitution $t = -t_2 + t'$ one obtains:

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_2 - \frac{T}{2}}^{t_2 + \frac{T}{2}} \xi^{(k)}(t_1 - t_2 + t') \,.\, \xi^{(k)}(t')\, dt' = R_{\xi\xi}^{(k)}(t_1 - t_2)$$

and, hence, $R_{\xi\xi}^{(k)}(t_2-t_1) = R_{\xi\xi}^{(k)}(t_1-t_2)$. The autocorrelation, therefore, is an *even function*.

4. **Cross-correlation function**:

$$R_{\xi\eta}^{(k)}(t_1 - t_2) \triangleq \widetilde{\xi^{(k)}(t_1 + t) \,.\, \eta^{(k)}(t_2 + t)}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \xi^{(k)}(t_1 + t) \,.\, \eta^{(k)}(t_2 + t)\, dt = R_{\xi\eta}^{(k)}(t_2 - t_1)$$

1.8 CLASSIFICATION OF STOCHASTIC SIGNALS

1.8.1 MIXED SIGNALS

In many practical situations a deterministic signal s(t) is considered, to which a stochastic signal $\xi(t)$ is added; e.g. additive noise that is superpose with an information bearing signal in a communication channel. Such signals are labelled *mixed* signals. So:

$$\eta(t_1) = \xi(t_1) + s(t_1)$$

Let $w_1(x_1;t_1)$ be the probability density function of the stochastic signal $\xi(t_1)$ and $W_1(y_1;t_1)$ the one of the stochastic signal $\eta(t_1)$. What would then be the relationship between the densities w_1 and W_1?

$W_1(y_1;t_1)dy = w_1(x_1;t_1)dx$ since $s(t_1)$ = a constant. However, $w_1(x_1;t_1)dx$ is the probability that $\xi(t_1)$ would attain values between x and x + dx. Taking into account that y = x + $s(t_1)$, this indicates that dy = dx, and therefore $W_1(y_1;t_1) = w_1(x_1;t_1)$; or

$$W_1(y_1;t_1) = w_1[y_1 - s(t_1);t_1] \qquad (\ddagger)$$

1.8.2 WIDE SENSE STATIONARY SIGNALS

When the statistical properties of signals are invariant of any arbitrary shift in time, then the signals are considered to be *wide sense stationary*. In other words:

$$w_n(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n) = w_n(x_1,\ldots,x_n;t_1+\tau,\ldots,t_n+\tau)$$

In case that this property is valid $\forall\tau$, then the signal is called *wide sense stationary*.

Consider as an example the signal $\eta(t) = \cos(\omega_0 t+\zeta)$, where has $\eta(t)$ a probability density function $w_1(x)$ in the interval $[0.2\pi]$. To check whether the signal $\eta(t)$ is stationary, one will shift it with τ with respect to the time origin.

$\eta(t+\tau) = \cos(\omega_0 t+\omega_0 \tau+\xi) = \cos(\omega_0 t+\zeta)$, where $\zeta = \xi+\omega_0 \tau$.

By definition, ζ is a *mixed signal*, since it exists of the sum of ξ, a *stochastic process*, and $\omega_0 \tau$, a *deterministic* component.

From ($\ddagger$) the probability density can be derived: $\eta(t+\tau) = w_1(y-\omega_0 \tau)$. To ensure that the signal $\eta(t)$ is stationary, the density $W_1(y;t+\tau)$ should be independent of τ. This is only the case if $w_1(y-\omega_0 \tau)$ is independent of τ, or in other words, if the stochastic variable ξ has a uniform distribution in the interval $[0,2\pi]$. This is achieved if $w_1(x)=1/2\pi$. In Figure 1.3 this conclusion is illustrated.

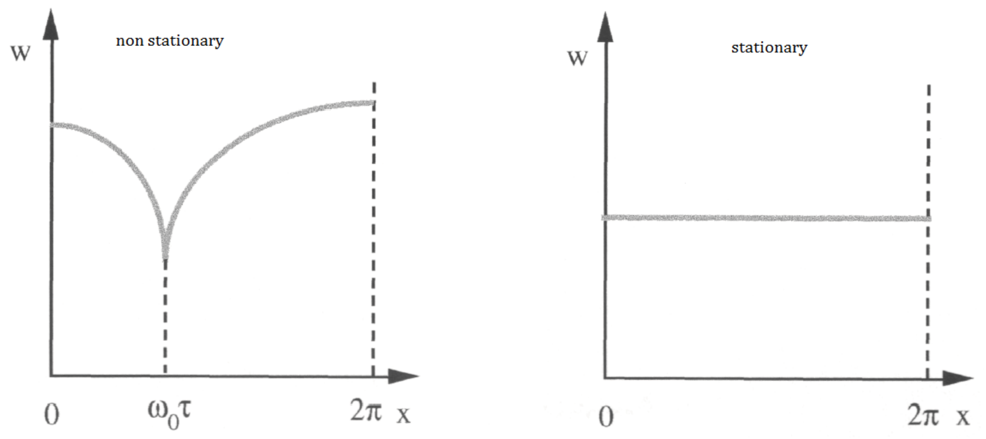

Figure 1.3: Examples for the p.d.f. w(x) of a stochastic process ξ.

The condition for wide sense stationarity for a signal:

$$w_n(x_1,...,x_n;t_1,...,t_n) = w_n(x_1,...,x_n;t_1 + \tau,...,t_n + \tau)$$

must be valid $\forall\tau$, and hence also for $\tau=-t_1$, so that this condition can be rewritten as:

$$w_n(x_1,...,x_n;t_1,...,t_n) = w_n(x_1,...,x_n;0,t_2 - t_1,...,t_n - t_1)$$

When n=1, this results in:

$$w_1(x_1;t_1) = w_1(x_1)$$

Hence, the density of first order is independent from the time. When n=2 one obtains:

$$w_2(x_1,x_2;t_1,t_2) = w_2(x_1,x_2;t_2 - t_1)$$

The density of order two is depending only on the time interval t_2-t_1.

Substituting of these conditions into the definitions of the ensemble averages one obtains:

$$\overline{\xi(t_1)} = \int_{-\infty}^{+\infty} x_1 \cdot w_1(x_1)\, dx_1 = \mu = C^{te} \quad \forall\, t_1!$$

The ensemble mean is independent of the time.

The variance becomes:

$$\sigma^2(t_1) \triangleq \overline{\left[\xi(t_1) - \overline{\xi(t_1)}\right]^2} = \int_{-\infty}^{+\infty} (x - \mu)^2\, w_1(x)\, dx = \sigma^2 = C^{te}$$

Hence, also the variance is independent of the time.

With respect to the autocorrelation function one will observe the dependency on the time interval t_2-t_1 only. This property is always valid for the autocorrelation function of the realization. Only wide sense stationary signals can claim the similar result for the autocorrelation function of the ensemble.

$$B_{\xi\xi}(t_1,t_2) = \overline{\xi(t_1)\,.\,\xi(t_2)} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1\, x_2\, w_2(x_1,x_2;t_2-t_1)\, dx_1 dx_2$$

$$= B_{\xi\xi}(t_2 - t_1)$$

1.8.3 ERGODIC SIGNALS

Signals are *ergodic* when the ensemble averages and the averages of the realizations are equal to each other.

$$\overline{\xi(t)} = \widetilde{\xi^{(k)}(t)}$$

When one single realization is observed for the signal, only the realization averages can be calculated. For ergodic signals from a single realization the complete set of statistical properties can be derived.

It is easy to prove that a condition ensuring that signals are ergodic is that they should be wide sense stationary. Unfortunately this condition is to be required only, but, unfortunately, not sufficient.

An example will illustrate this. Let $\eta(t) = \cos(\xi t+\varphi)$, a stationary signal, where ξ and φ are independent stochastic variables. ξ has an arbitrary distribution w(x) and φ a uniform one in the interval $[0,2\pi]$ (otherwise the signal will not be wide sense stationary). The statistical ensemble mean is:

$$\overline{\eta(t)} = \int_{-\infty}^{+\infty}\int_{0}^{2\pi} \left[\cos(\xi t).\cos(\varphi) - \sin(\xi t).\sin(\varphi) \right] \,.\, w(\xi)\,.\,\frac{1}{2\pi}\, d\xi d\varphi$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos(\xi t)\,.\,w(\xi)\, d\xi \underbrace{\int_{0}^{2\pi} \cos(\varphi)\, d\varphi}_{\Downarrow\ 0} - \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sin(\xi t).w(\xi)\, d\xi \underbrace{\int_{0}^{2\pi} \sin(\varphi)\, d\varphi}_{\Downarrow\ 0}$$

$$\Rightarrow \quad \overline{\eta(t)} = 0$$

The realization average yields:

$$\widetilde{\eta(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} \cos\left[\xi^{(k)}(t).t + \phi^{(k)} \right] dt = 0$$

The evolution of the signal η(t) is illustrated in Figure 1.4.

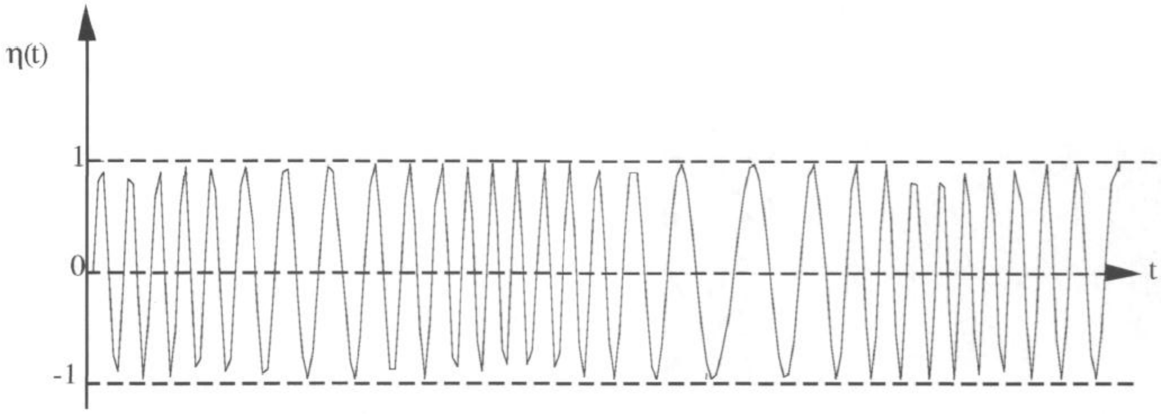

Figure 1.4: The evolution of the signal η(t) for the realization k.

Hence:

$$\overline{\eta(t)} = \widetilde{\eta(t)}$$

This process, therefore, is ergodic.

Consider now, however, the signal η(t) = ξ(t) + ζ, where ξ(t) is a stationary ergodic process and ζ a stochastic variable. The process is stationary, but not ergodic.

$$\overline{\eta(t)} = \overline{\xi(t)} + \overline{\zeta} = x_0 + z_0$$

Where x_0 and z_0 are the ensemble averages, and:

$$\widetilde{\eta(t)} = \widetilde{\xi(t)} + \zeta = x_0 + \zeta$$

In other words: the process is only ergodic if ζ is a constant.

1.8.4 TOTALLY STOCHASTIC SIGNALS

Totally or *absolutely stochastic signals* are the simplest signals for analysis in the wide family of stochastic signals. Consecutive values of ξ are totally independent from each other; i.e. they are uncorrelated. In that case the values of the signal, evaluated in arbitrarily small intervals, are independent from each other. This can be expressed as:

$$w_n(x_n;t_n/x_1,x_2,\ldots,x_{n-1};t_1,\ldots,t_{n-1}) = w_1(x_n,t_n)$$

or:

$$w_n(x_n;t_n/x_1,x_2,\ldots,x_{n-1};t_1,t_2,\ldots,t_{n-1})dx_n$$

is the probability that $\xi(t)$ will take values between x_n and x_n+dx_n at the time instance t_n, when ξ has reached the values $x_1,x_2,\ldots,x_{n-1}$ at the time instances $t_1,t_2,\ldots,t_{n-1}$ where $t_1\leq t_2\leq\ldots\leq t_{n-1}$ and this probability will depend only on x_n at t_n and not from all previously attained values at the previous time instances.

For n=2 this is written as: $w_2(x_2;t_2/x_1;t_1) = w_1(x_2;t_2)$.

From the definition of the conditional probability one retains that:

$$w_2(x_2;t_2/x_1;t_1) \triangleq \frac{w_2(x_1,x_2;t_1,t_2)}{w_1(x_1;t_1)}$$

Hence, one can express that: $w_2(x_1,x_2;t_1,t_2) = w_1(x_1;t_1).w_1(x_2;t_2)$ and hence $\forall n$:

$$w_n(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n) = \prod_{k=1}^{n} w_1(x_k;t_k)$$

As such, the values $x_1,x_2,\ldots,x_n$ are indeed independent from each other $\forall t_k \neq t_j$. This relation also clearly states that totally stochastic signals are entirely described when the density of the first order is known, since all higher orders can be derived from the latter.

1.8.5 MARKOVIAN SIGNALS

The stochastic signals, which can be described with the knowledge of the first and second order statistical moments only, are labelled *Markovian* signals. They are important in many telecommunication studies. In first order Markov processes the value x_n at the time instance t_n will depend on the previous values $(x_1,x_2,\ldots,x_{n-1})$ at the time instances $(t_1,t_2,\ldots,t_{n-1})$ via the last obtained value $(x_{n-1};t_{n-1})$. In other words:

$$w_n(x_n;t_n/x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1}) = w_2(x_n;t_n/x_{n-1};t_{n-1})$$

In general one can write that:

$$w_n(x_n,t_n/x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1}) = \frac{w_n(x_1,\ldots,x_n;t_1,\ldots,t_n)}{w_{n-1}(x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1})}$$

so that:

$$w_n(x_1,\ldots,x_n;t_1,\ldots,t_n) = w_{n-1}(x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1}) \cdot w_2(x_n;t_n/x_{n-1};t_{n-1})$$

Accordingly, one can write that:

$$w_{n-1}(x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1}) = w_{n-2}(x_1,\ldots,x_{n-2};t_1,\ldots,t_{n-2}) \cdot w_2(x_{n-1};t_{n-1}/x_{n-2};t_{n-2})$$

$$\ldots$$

$$w_3(x_1,x_2,x_3;t_1,t_2,t_3) = w_2(x_1,x_2;t_1,t_2) \cdot w_2(x_3;t_3/x_2;t_2)$$

$$w_2(x_1,x_2;t_1,t_2) = w_1(x_1;t_1) \cdot w_2(x_2;t_2/x_1;t_1)$$

and, hence:

$$w_n(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n) = w_1(x_1;t_1) \cdot \prod_{k=2}^{n} w_2(x_k;t_k/x_{k-1};t_{k-1})$$

One will agree that the process is completely described, when the statistical moments up to order two are known.

On the other hand, one will conclude that the values $(x_1,x_2,\ldots,x_{n-2})$ at the time instances $(t_1,t_2,\ldots,t_{n-2})$ are of no importance to determine the value x_n at the time t_n. One can express this in other words stating that the value of the future sample x_n will not depend on the values obtained in the past $(x_1,x_2,\ldots,x_{n-2})$, but only on the actual one x_{n-1}.